STAT 610: Discussion 1

1 Summary

- Order statistic: Let $X_1, ..., X_n$ be i.i.d sample and $X_{(1)}, ..., X_{(n)}$ as order statistics, then

$$F_{X_{(j)}}(x) = \sum_{k=j}^{n} \binom{n}{k} \{F_X(x)\}^k \{1 - F_X(x)\}^{n-k}.$$

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!} f_X(x)\{F_X(x)\}^{j-1}\{1 - F_X(x)\}^{n-j}.$$

$$\begin{aligned} f_{X_{(i)},X_{(j)}}(u,v) = & \frac{n!}{(i-1)!(j-1-i)!(n-j)!} f_X(u) f_X(v) \{F_X(u)\}^{i-1} \\ & \times \{F_X(v) - F_X(u)\}^{j-1-i}\{1 - F_X(v)\}^{n-j} I(u < v). \end{aligned}$$

$$f_{X_{(1)},\ldots X_{(n)}}(x_1, \ldots x_n) = \begin{cases} n! f_X(x_1) \cdots f_X(x_n) & \text{if } -\infty < x_1 < \cdots < x_n < \infty, \\ 0 & \text{otherwise.} \end{cases}$$

- Sufficient statistics
  - Definition: A statistic $T(\mathbf{X})$ is sufficient for $\boldsymbol{\theta}$ if the conditional distribution of the sample $\mathbf{X}$ given the value of $T(\mathbf{X})$ does not depend on $\boldsymbol{\theta}$.
  - Factorization Theorem: $T(\mathbf{X})$ is sufficient for $\boldsymbol{\theta}$ if and only if

$$f(\mathbf{x}|\boldsymbol{\theta}) = g(T(\mathbf{x})|\boldsymbol{\theta})h(\mathbf{x}).$$

- Rao-Blackwell Theorem: Let X be a sample from a population indexed by $\theta \in \theta$ and statistic $T(\mathbf{X})$ is sufficient for θ. If $U(X)$ is a statistic used to estimate $\vartheta = \phi(\theta)$, and $E_\theta[U(X) - \vartheta]^2 < \infty$, then the statistics $h(T) = E[U(X)|T]$ satisfies

$$E_\theta[h(T) - \vartheta]^2 < E_\theta[U(X) - \vartheta]^2 \qquad \theta \in \theta$$

  and

$$E_\theta[h(T) - \vartheta]^2 = E_\theta[U(X) - \vartheta]^2 \qquad \theta \in \theta$$

  if and only if $P_\theta(U(X) = h(T)) = 1$, $\theta \in \theta$.

  - Sufficiency make sure h is a statistics.

2 Questions

1. Let $X_1, ..., X_n$ be a random sample from a population with pdf

$$f_X(x) = \begin{cases} \frac{1}{\theta} & \text{if } 0 < x < \theta \\ 0 & \text{otherwise.} \end{cases}$$

   Let $X_{(1)} < ... < X_{(n)}$ be the order statistics. Show that $\frac{X_{(1)}}{X_{(n)}}$ and $X_{(n)}$ are independent.

2. Show that if T is sufficient statistic and $T = \phi(S)$, where ϕ is a real function and S is another statistic, then S is sufficient.

3. Let $X_1 \ldots, X_n$ be i.i.d. random variables from the exponential distribution $E(a, \theta)$(i.e. $f_{a,\theta}(x) = \theta^{-1} e^{-(x-a)/\theta} I_{x>a}$). Find a two-dimensional sufficient statistics for (a, θ)

4. Let X and Y be two random variables such that Y has the binomial distribution with size N and probability π and, given $Y = y$, X has the binomial distribution with size y and probability p.

   (a) Suppose that π and N are known and $p \in (0, 1)$ is unknown. Show whether X is sufficient for p and whether Y is sufficient for p.